Running head: Time Course of Written Word Production

An Electrophysiological Analysis of the Time Course of Phonological and Orthographic Encoding in Written Word Production

Qingqing Qu[1 2] and Markus F. Damian[3]

[1]Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China

[2]Department of Psychology, University of Chinese Academy of Sciences, Beijing, China

[3]School of Psychological Science, University of Bristol, United Kingdom

Address for correspondence:

Qingqing Qu
Key Laboratory of Behavioral Science
Institute of Psychology, Chinese Academy of Sciences, Beijing,
16 Lincui Road, Chaoyang District, Beijing, 100101
China
Tel: +86-10-64888629
Fax: +86-10-64872010
Email: quqq@psych.ac.cn

Abstract

Recent evidence suggests that individuals generate written words based on both spelling and sound. The present study used event-related potentials (ERPs) to examine the relative time course of orthographic and phonological activation. We adopted Chinese as a target language in which spelling and sound are largely dissociated. Native speakers of Chinese Mandarin were presented with colored pictures and wrote down color and picture names as adjective-noun phrases. Color and picture names were either phonologically related, orthographically related, or unrelated. EEG revealed phonological effects in the 200-500 ms time window, starting at 206 ms after picture onset, and orthographic effects in the 300-400 ms time window, starting at 298 ms. The results of our study suggest that activation of phonological codes takes place approximately 100 ms earlier than access to orthographic codes, which provides evidence for phonological encoding as early sources of constraint in written word production.

Key words: Handwriting; written production; orthography; phonology; electrophysiology; Chinese

An Electrophysiological Analysis of the Time Course of Phonological and Orthographic Encoding in Written Word Production

Introduction

Over the past few decades, a vast amount of psycholinguistic research has been dedicated to exploring how individuals comprehend and produce spoken language, as well as how they process written (orthographic) codes. By contrast, comparatively little work has focused on tasks that require the generation of orthographic codes, such as writing and typing. In the work reported here, we focused on the handwriting of short utterances consisting of two words. We specifically investigated the relative time course of two cognitive processes underlying written production, i.e., phonological and orthographic encoding.

In research on written output processes, a central theoretical issue is whether written word production is constrained by phonological codes. According to a "phonological mediation" view advocated by early theorists (e.g., Geschwind, 1969; Luria, 1970), access to orthographic codes is based exclusively on prior retrieval of phonological codes. In other words, writing is essentially based on "inner speech", with subsequent translation of phonological into orthographic codes. Although plausible in the sense that spoken language precedes written language both ontogenetically and phylogenetically (e.g., Scinto, 1986), this position is no longer tenable because studies of individuals with acquired brain damage have demonstrated a dissociation between spoken and written production, e.g., patients who are unable to name pictures due to a deficit in the phonological lexicon, yet with preserved ability to write down picture names (e.g., Bub & Kertesz, 1982). Hence, the current consensus is reflected in the "orthographic autonomy" view, according to which orthographic codes can be accessed directly from meaning (Rapp, Benzing, & Caramazza, 1997).

A direct processing pathway from meaning to orthography does not preclude a role of phonology in written production, however. Indeed, a growing number of behavioral studies have documented phonological effects in a range of experimental tasks requiring handwritten output, such as the simple object naming (Bonin, Peereman & Fayol, 2001), "implicit priming" (Afonso & Álvarez, 2011), long-term priming (Damian, Dorjee & Stadthagen-Gonzalez, 2011), picture-word interference (Qu, Damian, Zhang, & Zhu, 2011; Zhang & Damian, 2010), a written Stroop (Damian & Qu, 2013), and masked priming (Qu, Damian & Li, 2015) tasks, although it is noted that effects of phonology failed to emerge in a few studies (Bonin, Fayol & Peereman, 1998; Shen, Damian & Stadthagen-Gonzalez, 2013). Because in alphabetic writing systems, orthography and phonology are necessarily confounded, it is difficult to design experiments to disentangle the two factors. A few studies have therefore recently begun to investigate handwriting in languages with non-alphabetic scripts such as Chinese, where sound and spelling are more easily dissociated from one another, and as in languages with alphabetic scripts, phonology appears to influence orthographic production (Damian & Qu, 2013; Qu, Damian, Zhang, & Zhu, 2011; Qu, Damian & Li, 2015). Overall, we perceive as the current consensus that the orthographic codes on which written word production is based are accessed not only via a direct link from semantics, but also via an indirect route via phonology. The claim that sound influences written production matches well with the everyday observation that writing or typing errors often result in homophones or pseudohomophones (e.g., producing “there” when intending to write/type “their”; producing “智力” /zhi4li4/, “intelligence”, when intending to write/type “智利”, /zhi4li4/, “Chile”).

Whether the proposed phonological route operates at a lexical or sub-lexical level is less well understood. Bonin, Peereman and Fayol (2001) investigated written word production in French and manipulated sound-to-print consistency of picture names. They found that when inconsistency was defined at the lexical level, writing latencies for heterographic homophones (e.g., "verre" and "vert") did not differ from nonhomophone controls. When inconsistency was defined sublexically, the position of the inconsistency within the target word was critical, with only word-initial, but not medial or final positions, affecting latencies. This pattern suggests that information is transmitted from phonology to orthography via a sequentially operating sublexical conversion route, rather than via lexical mappings. By contrast, Wang and Zhang (2015) investigated Chinese written word production and reported an inhibitory phonetic regularity effect (low-frequency picture names with regular first characters were slower to write than with irregular ones) as well as an inhibitory homophone density effect (characters with many homophones were produced more slowly than characters with few homophones). The authors concluded that for Chinese individuals, phonology influences orthographic word production both at a lexical and a sub-lexical level.

Further critical evidence with regard to the processing characteristics of handwriting comes from the study of event-related potentials (ERPs), which allow the tracking of neural responses associated with mental activities by the millisecond before a response is being executed. Recent ERP studies have explored in detail *spoken* word production (e.g., Costa, Strijkers, Martin, & Thierry, 2009; Qu, Damian, & Kazanina, 2012; Strijkers, Costa, & Thierry, 2010; Strijkers, Holcomb, & Costa, 2011; see Ganushchak, Christoffels, & Schiller, 2011, and Strijkers & Costa, 2011, for reviews). Based on these and related studies, the current time estimates of various cognitive stages involved in preparing a spoken picture naming response and based on an assumed naming latency of 600 ms are, conceptual preparation: 0-200 ms; lexical-semantic (“lemma”) retrieval: 200-275 ms; form encoding: 275-600 ms (Indefrey, 2011). Given the status of “lemmas” as abstract lexical entities (Levelt, Roelofs & Meyer, 1999) it is likely that these are shared across different output modalities such as speaking and writing (and perhaps even across production and reception tasks). By contrast, subsequent word form encoding should be modality-specific because spoken and written production involve very different output codes. The currently available evidence hence suggests the model outlined in Figure 1: spoken and written picture naming share processing components up to and including “lemma” access, then diverge into modality-specific components. Writing is carried out via a “direct access route" (pathway A in Figure 1). At the same time, activation propagates from the lemma level to the phonological level, and then impacts orthographic encoding via cross-talk (pathway B in Figure 1).

Critical questions arise regarding the temporal properties of access to these types of codes. Do both pathways deliver activation to the orthographic level at the same pace, or is one faster than the other? Because spoken language is acquired earlier in life and used more frequently than written language, it could be hypothesized that phonological codes are rapidly (and perhaps automatically, e.g., Bles & Jansma, 2008) accessed from meaning; by contrast, access to orthographic codes is more effortful, less rapid and more indirect. Conversely, one could argue that the ultimate purpose of writing is to express meaning, hence orthographic encoding should be mainly based on direct input from semantics, with additional - and perhaps delayed - activation delivered via the indirect phonological route.

It is difficult to identify the time course of orthographic vs. phonological encoding via behavioral measures alone. Nonetheless, a few studies have begun to explore this issue, mainly via use of the picture-word interference task in which individuals name, via speaking or writing, an object while attempting to ignore a "distractor" word (e.g., Lupker & Katz, 1981). In this task, the interval between picture and distractor onset (stimulus-onset asynchrony; SOA) can be varied, with the rationale that this allows the distractor to tap into successive stages of target processing, with negative SOAs affecting early stages of target processing, and positive SOAs influencing later stages (e.g., Schriefers, Meyer & Levelt, 1990). In versions of this task which required written picture naming, it was found both for English (Zhang & Damian, 2010) and for Chinese participants (Qu, Damian, Zhang, & Zhu, 2011; see also Damian & Qu, 2013 for relevant evidence from a Stroop task) that phonological effects emerged at a relatively "early" point in time (e.g., phonological effects were restricted to the 0-ms SOA in Qu, Damian, Zhang, & Zhu, 2011 and Zhang & Damian, 2010), whereas priming effects at a "later" SOA (e.g., +100-ms SOA) were exclusively determined by orthographic properties. This pattern potentially implies that the activation of phonological codes precedes the retrieval of orthographic codes. However, with the same experimental approach, Zhang and Wang (2015) more recently found the opposite pattern: an exclusively orthographic effect at an early stage (SOA = -100 ms), and orthographic and phonological effects at later stages (SOA = 0 ms and +100 ms). The discrepancy between these and the earlier findings is not well understood but should perhaps caution against drawing strong inferences from picture-word and related tasks regarding the underlying time course of cognitive processes.

Additional relevant evidence comes from a number of recent EEG studies. Perret and Laganaro (2012) conducted a study in which spoken and written responses to the same pictorial stimuli were directly compared. The authors found identical electrophysiological activity associated with written and spoken responses for the initial stages of picture naming. The two modalities began to diverge and display modality-specific characteristics beginning about 260 ms post picture onset, which roughly matches the time estimate for phonological encoding proposed by Indefrey (2011; Indefrey & Levelt, 2004). Subsequent to the 260 ms point in time, they identified two separate time windows (260-400 ms, and 400-600 ms), and attributed the earlier window to orthographic or phonological form encoding. Other relevant evidence comes from a picture naming EEG study in which Perret, Bonin, and Laganaro (2014) manipulated the age of acquisition (AoA) of picture names, a variable which is commonly assumed to reside at the level of word-form encoding. They found that the spatiotemporal maps of the late and early AoA conditions presented different time distributions beginning at 260 ms post picture onset (with a longer time interval for late-acquired words compared to early-acquired words), a pattern which suggests that word-form encoding in writing initiates around 260 ms. Other than AoA, in two recent studies word frequency was manipulated in orthographic production tasks. Baus, Strijkers, and Costa (2013) asked Spanish participants to type picture words, and found a relatively "late" effect of frequency in ERPs (i.e., 330-430 ms). By contrast, Qu, Zhang and Damian (2016) investigated frequency effects in the written production of Chinese words, and found that ERPs elicited by high- and low frequency items started to diverge as early as 168 ms post picture onset.

However, directly relevant evidence from ERP studies concerning the time course of orthographic and phonological encoding in written production is scarce. Zhang and Wang (2016) used a written picture-word interference task and manipulated orthographic and phonological overlap between picture names and distractor words. ERPs showed an early orthographic effect (370-500 ms) and a late phonological effect (460-500 ms). However, note that as outlined above, the behavioural results from picture-word tasks concerning this issue (via manipulation of target-distractor SOAs) are inconsistent. Moreover, concerns about the picture-word interference task for investigating the time course of production have been voiced (Strijkers & Costa, 2011). One possible criticism is that the exact locus of word-form effects from the picture-word interference task remains controversial. For example, orthographic effects have been attributed to an early, conceptual level (Zhang & Weekes, 2009), or alternatively, to a late, word-form encoding level (Zhao, La Heij, & Schiller, 2012). Moreover, in the picture–word interference task, the superimposition of a visually complex distractor could itself delay the cognitive processes associated with written target word production. Indeed, the time at which Zhang and Wang observed orthographic effects (370-500 ms) is comparably late given the estimates of when word-form encoding is expected to take place (i.e. beginning at 275 ms; Indefrey, 2011).

In the present study, we tackled the issue of the relative time course of orthographic vs phonological encoding in written production via measurement of EEG. We used a task in which on each trial, participants were presented with a line drawing of a common object, with the lines of the object colored, and were asked to write down the color and picture name as an adjective-noun phrase ("orange chair"). Previous studies on spoken production in which this task was used have shown that word-form overlap between the color and picture names (e.g., **g**reen **g**oat) accelerates response latencies, relative to a condition in which color and picture name are unrelated (Damian & Dumay, 2007, 2009). Similar facilitation is found when responses are written on a digital graphic tablet, rather than being spoken (Damian & Stadthagen-Gonzalez, 2009). In the current study, we recruited Chinese participants, and we independently manipulated orthographic and phonological overlap between adjective and noun. This is very difficult to accomplish in languages with alphabetic orthography, but achievable in a language with a non-alphabetic script such as Chinese Mandarin. Both phonological and orthographic overlap between the color and picture names was manipulated at the sublexical level: in the **phonologically** related condition, color and picture names shared a rhyme, but orthographic overlap was avoided. In the **orthographically** related condition, color and picture names shared an orthographic radical, but did not overlap in phonology (see Figure 2A). By measuring ERPs while participants prepared their written responses, we expected to identify separate modulations of ERPs associated with the manipulation of sound and spelling. Doing so should be informative with regard to the relative time course of access to phonological and orthographic representations in written preparation.

Method

Participants

Thirty native speakers of Mandarin Chinese (9 males, ages 19-26, mean age 21) from Beijing Forestry University and China Agricultural University participated in the experiment for monetary compensation. All were right-handed, had normal or corrected-to-normal vision. None were color blind or with any known language deficit. Four of the participants were subsequently removed from the analyses (behavioral and EEG analyses), due to a high number of rejected trials (more than 20%, range: 20.3%-34.9%). Statistical analyses are thus based on 26 individual data sets (rejected trials: 2.6%-17.7%).

Materials and Design

Four colors (red, blue, brown, and orange) were used. The corresponding color names in Chinese are monosyllabic. Twenty-four line drawings of pictures with no canonical color were chosen from the Snodgrass and Vanderwart (1980) picture set, with the majority of corresponding picture names being disyllabic (see the *Appendix* for a complete list of experimental materials). Note that as in English, adjectives precede nouns in Chinese. In the *phonologically related* (P) condition, each color was combined with three pictures to form 12 phonologically related combinations in which the picture and color name shared a *rhyme* (e.g., 橙灯泡, /cheng2deng1pao4/, 'orange lightbulb'). In the *orthographically related* (O) condition, each color was combined with three pictures to form 12 orthographically related combinations in which both shared a *radical* in the same positions within a character (e.g., 橙椅子, /cheng2yi3zi/, 'orange chair'). Because the size of effects depends on the degree of overlap, therefore, we manipulated both the orthographic and the phonological relatedness at the sub-character level to keep the degree of overlap as comparable as possible across the P and O conditions. Moreover, across the P and O conditions, pictures were statistically matched on various lexical properties.[1]To form the respective *unrelated* conditions, colors and pictures were recombined to avoid orthographic and phonological overlap between the picture and color name. Hence, each participant was presented with four blocks of 48 trials with each of the 24 related and 24 unrelated color-picture combinations appearing exactly once in each block, for a total of 192 trials (48 trials in each of four conditions). A new pseudorandom order was generated for each block and participant. Neither pictures nor colors were repeated on consecutive trials.

Procedure

Stimuli were presented using E-Prime 1.1 software (Psychology Software Tools, Pittsburgh, PA). Written responses were recorded using an Intuos4 graphic tablet and inking pen (Wacom, Kazo-shi, Japan). Participants were tested individually in a sound-attenuated lab. Participants were first asked to familiarize themselves with the experimental stimuli by viewing them in a booklet, with the expected name printed underneath each picture. Subsequently, participants were told that they would see the pictures in different colors presented in the center of the computer screen, and their task was to write down color and picture name with an adjective-noun combination as quickly and accurately as possible, e.g., 棕枕头, /zong1zhen3tou2/, 'brown pillow'. Participants were instructed to lift the pen very slightly from the answering sheet so that response could be given as fast as possible; they should not drop the pen on the sheet before identifying the response. They were asked to keep gazing at the screen during writing and refrain from looking at what they had written (i.e., visual feedback was prevented) in order to minimize movement artifacts in the EEG recording. Compliance with these instructions was assured before the experiment began.

In a subsequent practice block, 12 pictures with colors which were not related to the picture names were presented. After the practice, four experimental blocks of 48 trials were presented. There were short breaks between blocks, and the next block started after participants indicated that they were ready to continue. On each trial, participants saw a sequence consisting of a fixation cross (500 ms), a blank screen (500 ms), and a picture. The picture disappeared once the participant initiated a response on the graphic tablet, or after a time-out of 4,000 ms. The intertrial interval was 6,000 ms. The experimental task session lasted approximately 40 minutes. The entire experiment lasted about 2 hours.

EEG recordings

---

[1]Stimuli in the phonological and orthographic conditions were matched on the following variables: number of strokes, word frequency, word length in number of character, image variability, image agreement, concept familiarity, visual complexity, subjective frequency, name agreement, concept agreement, and age-of-acquisition. Values were taken from Liu, Hao, Li, and Shu (2011).

The electroencephalogram (EEG) was recorded with 64 electrodes secured in an elastic cap (Electro Cap International) using Neuroscan 4.3 software. The vertical electrooculogram (VEOG) was monitored with electrodes placed above and below the left eye. The horizontal EOG (HEOG) was recorded by a bipolar montage using two electrodes placed on the right and left external cantus. The left mastoid electrode served as reference. The EEG data were re-referenced off-line to the average of both mastoids. All electrode impedances were kept below 5 kΏ. Electrophysiological signals were amplified with a band-pass filter of 0.05 and 100 Hz (sampling rate 500 Hz) and filtered off-line using a 30 Hz low-pass (zero-phase) filter.

Recordings were analyzed offline using Neuroscan 4.3 software. The VEOG electrode activity was applied to ocular artifact correction; the ocular artifact was conducted with a negative-going EEG at 10% with 40 minimum sweeps with durations of 400 ms. Epochs containing artifact signals below/above ±75µV were rejected. The 500 ms post-stimulus interval was chosen in order to guarantee covering the time window of core processes of word production (Indefrey & Levelt, 2004) and minimize contamination of EEG signals from movement due to writing. The EEG was segmented into 600 ms epochs relative to picture onset that included a 100 ms pre-stimulus interval and a 500 ms post-stimulus interval.[2]

Data analysis

Trials with incorrect responses (2.9%) and trials with naming onset latencies faster than 500 ms or slower than 2,000 (5.5%) were excluded from the behavioral and ERP analyses. For the ERP analysis, a further 1.0% of trials were excluded due to artefacts. In total, ERP analyses were based on an average of 43.5 segments per condition (orthographically related: 44, orthographically unrelated: 43, phonologically related: 44, phonologically unrelated: 43).

Two types of analyses were conducted on the ERP data. First, onset latency analysis was performed, with the aim of identifying the latency at which the ERPs of critical conditions started to diverge from each other (orthographically related vs. unrelated, and phonologically related vs. unrelated). To protect against problems associated with multiple comparisons, we performed onset latency analyses using a method developed by Guthrie and Buchwald (1991) (see e.g., Costa, Strijkers, Martin, & Thierry, 2009; Qu, Zhang, & Damian, 2015; Strijkers, Costa, & Thierry, 2010; Thierry, Cardebat, & Demonet, 2003 for use of this method in recent studies). This method assumes difference waveforms possess a first-order autoregressive structure with sampling points statistically dependent and uses this assumption to generate how long an interval of consecutive significant points can be expected by chance (i.e., "the critical run length for determining

[2]Laganaro and Perret (2011) and Perret and Laganaro (2012) introduced an analysis which combines stimulus- and response-aligned ERPs in production tasks. This form of analysis is undeniably valuable because almost the entire production process can be covered. However, in our study, EEG signals close to the average written latency were substantially contaminated from movement due to writing; therefore, we exclusively performed stimulus-aligned analysis.

statistical significance", the duration of interval that you count as significant) via computer simulations. Computer-simulated estimates of such critical run length were based on 1000 repetitions for each of several autocorrelation coefficients, sample sizes, and sampling interval length. If the observed number of consecutive significant time points is larger than the critical run length, it would indicate a statistically significant interval. The onset point of a sequence of consecutive significant points is deemed as the onset of the effects.

Second, mean amplitudes analyses were conducted. Mean amplitudes were calculated separately for each participant and each condition in each time windows. With the combined consideration of results of onset latency analyses[3] and visual inspection of the ERP waveforms, the following four time windows were selected for statistical analyses: 0-200, 200-300, 300-400, 400-500 ms. To provide a comprehensive picture of ERP effects, we conducted statistical analyses using six regions of interest (ROIs), each representing an average of 3 electrodes: left-anterior (electrodes: F5, F7, FC5), mid-anterior (Fz, FCz, Cz), right-anterior (F6, F8, FC6), left-posterior (P5, P7, CP5), mid-posterior (CPz, Pz, POz), and right-posterior (P6, P8, CP6). In this ROI analysis that enabled us to probe the scalp distribution of ERP differences, mean amplitudes from each time window were entered into a 2×2×6 repeated measures ANOVA with type (phonological/orthographic), relatedness (related/unrelated), and regions. Greenhouse-Geisser correction was applied where appropriate, to control for violations of the sphericity assumption (original degrees of freedom are reported). Effects for the electrode region are only reported when they interact with the experimental manipulations.

Results

Behavioral data

Behavioral data are summarized in Figure 2B. Response latencies were analyzed with a linear mixed-effects model using the software R (R Development Core Team, 2009) with the package lme4 (Bates & Maechler, 2009) with participants and items as random factors and type (phonological/orthographic) and relatedness (related vs. unrelated) as fixed factors. Model fitting was carried out by initially specifying a model that only included the random effects (participants and items) which was then enriched by adding the fixed factor type, followed by relatedness, and finally the interaction between the two factors.[4] The best fitting model was defined to be the most complex model that significantly improved the fit over the previous

[3]The onset latency analysis suggests ‘‘breakpoints" around 200 ms (where phonological effects begin to appear) and 300 ms (where orthographic effects begin to appear), around 400 ms (where orthographic effects disappear) and around 500 ms (where phonological effects disappear).

[4]For both latencies and errors, we conducted additional analyses in which the fixed factor repetition was included, and we obtained a main effect of repetition ($p < .01$, average latencies accelerated with repeated naming of the same trials). But critically, repetition did not statistically interact with any of the other factors, *p*s > .56".

model.[5] The results revealed that adding the factor type did not improve the fit of model, χ^2s < 1.46, t = 1.20, p = .226. Adding the factor relatedness and the interaction between relatedness and type improved the fit of model (χ^2s > 4.36, $|t|$s > 2.4, ps < .015). Further analyses for each type separately revealed a null effect for the phonological condition (779 vs. 778 ms, $\chi^2 < 1$, $t < 1$, p = .810) and a robust effect for the orthographic condition (746 ms vs. 770 ms, $\chi^2 = 10.8$, t = 3.3, p = .001). A parallel analysis of variance conducted on the errors showed that none of the models including relatedness, type, or their interaction significantly improved the fit, $\chi^2 < 1$, Zs < 1.22, ps > .220.

EEG data

Onset latency analyses. As shown in Figure 2, ERPs displayed a typical P1-N1-P2 peak sequence classically observed for visual stimulus presentation. ERPs for orthographically related vs. unrelated, and phonologically related vs. unrelated were compared by running t-tests at every sampling point (every 2 ms) starting from picture onset (0 ms) until 500 ms over all 62 electrodes. Onset latency was computed on averages of those electrodes in which the observed number of consecutive significant time points was larger than the critical run length in order to determine statistical significance. For the phonological effect, the averaged splitting point computed from individual splitting point estimates (23 electrodes) was **206** ms after picture onset. The averaged splitting point for the orthographic conditions (25 electrodes) was **298** ms after picture onset.

Mean amplitude analyses. Grand average ERP waveforms are displayed in Figure 2C, 2D for the two conditions and for six regions of interests chosen for the analysis. The main results of the omnibus ANOVA, conducted separately for each of four 100 ms time intervals, as summarized in Table 1. In the **0-200 ms time window**, omnibus ANOVA showed that there were no main effects or interactions involving relatedness or type (all Fs < 1.35, ps > .250), except for the type × region interaction ($F(5, 125) = 3.29$, p = .008). In the **200-300 ms time window**, the type × region interaction was significant, $F(5, 125) = 6.53$, p = .001. Critically, there was an interaction of type and relatedness, $F(1, 25) = 6.87$, p = .015. Further analyses examining the types phonology and orthography separately revealed a significant interaction of relatedness and region for the phonology condition ($F(5, 125) = 2.84$, p = .045), reflecting a phonological effect in mid-posterior, $t(25)$ = -2.52, p =.018, and right-posterior regions, $t(25)$ = -2.58, p = .008. In contrast, there was no main effect or interactions with regions for the orthography condition (Fs < 1.93, ps > .178). In the **300-400 ms time**

[5]According to the argument highlighted in Barr, Levy, Scheepers and Tily (2013) one should specify a "maximum model" by including not only by-participant and by-item adjustments, but also allow for adjustments to the slope of each critical within-participants/items variable. Because both relatedness and type of relatedness are manipulated within-participants and with-items, we specified slope adjustments for participants and items. However, this model returned a correlation of 1.00 between intercept and slope for the critical variable, which indicates that the model has been overparameterized (Baayen, Davidson & Bates, 2008) and the simpler model without slope adjustments is preferable.

window, the type × relatedness, type × region, type × relatedness × region interactions were significant (*F*s > 3.35, *p*s < .017). Further analyses examining the types phonology and orthography separately revealed a significant main effect for the orthography condition with no significant interaction with regions, $F(1, 25) = 4.63$, $p = .041$, and a significant interaction of relatedness and region for the phonology condition, reflecting phonological effects in the mid-posterior ($t(25) = -2.47$, $p = .021$) and right-posterior regions ($t(25) = -2.73$, $p = .011$). In the **400-500 ms time window,** the main effect of type, type × relatedness, type ×region, type × relatedness × region interactions were significant (*F*s > 2.71, *p*s < .044). Follow-up analyses demonstrated that for the phonology condition, the relatedness × region interaction were significant, $F(5, 125) = 3.01$, $p = .031$. The interaction reflected a significant phonological relatedness effect at the mid-posterior ($t(25) = -2.75$, $p = .011$) and right-posterior regions ($t(25) = -2.58$, $p = .016$). For the orthography condition, none of effects involving the factor relatedness was significant (*F*s < 1.46, *p*s > .238).

Table 1. Analysis of Variance for mean amplitude with type, relatedness and region in the four time windows.

	F			
	0 -200	200 -300	300 -400	400 -500
type	0.39	0.93	2.22	13.44**
relatedness	0.32	0.03	0.08	0.52
type × relatedness	0.79	6.87*	7.39*	5.12*
type × region	3.29**	6.53**	15.02***	12.28***
relatedness × region	1.35	1.70	2.28	1.07
type × relatedness × region	0.15	2.40	3.54*	2.72*
Phonology				
relatedness	1.16	2.89	2.52	4.20
relatedness × region	0.68	2.84*	4.13**	3.01*
Orthography				
relatedness	0.02	1.96	4.63*	1.46
relatedness × region	0.60	1.35	1.42	0.56

Note. * $p < .05$; ** $p < .01$; *** $p < .001$.

Discussion

The aim of the present study was to investigate the time course of the activation of phonological and orthographic codes in written word production using an ERP technique with high resolution of temporal properties. Adopting a task in which Chinese participants were presented with colored pictures and wrote down colors and picture names as adjective-noun phrases, we manipulated orthographic and phonological relatedness between color and picture name. Before any further discussion, it is worthy to note that response latencies were similar for orthographic and phonological conditions (in the unrelated condition, 770 vs. 778, $t = 0.33$, $p = .750$; in the related condition, 746 vs. 779 ms, $t = 1.32$, $p = .20$), thus, it is unlikely that any difference in the time course of phonological and orthographic effects is due to the difference in response latencies. ERP data showed that phonological relatedness modulated ERPs in 200-500 ms time window broadly across posterior regions, whereas orthographic relatedness modulated ERPs in a 300-400 ms time interval following picture onset. Precise temporal analysis revealed that the phonological relatedness effect emerged with an onset of 206 ms post picture onset, whereas the orthographic effect had an onset of 298 ms. Hence, in written word production, activation of phonological codes precedes access to orthographic codes by approximately 100 ms.

The finding that ERP amplitudes were affected by phonological manipulations as early as 200 ms after picture onset is somewhat surprising, given the estimates for the time course of phonological encoding in spoken production by Indefrey (2011; Indefrey & Levelt, 2004) in which 200 ms is linked to lexical selection rather than to phonological encoding. In the light of differences across studies (mainly, the estimates are for bare noun production whereas our study involved the production of adjective-noun phrases), linking the estimates by Indefrey and colleagues to our observed temporal findings is questionable and should be taken with caution (see Strijkers & Costa, 2016 for a review). The early activation of word forms here is in fact consistent with recent findings in the spoken modality. Qu, Damian and Kazanina (2012) found that phonological manipulations modulated ERPs starting from 200 ms after picture presentation. Miozzo, Pulvermüller and Hauk (2014) used a multiple linear regression approach to MEG analysis and found two early and simultaneous effects within 200 ms of variables related to semantic and phonological processing. Similarly, Strijkers, Costa and Pulvermüller (2017) found early simultaneous activation of lexical and phonological word properties in picture naming in a MEG study. Similar to these findings in the spoken modality, the results in the present study with written production suggest an early activation of phonological codes, which fits with recent accounts of parallel activation of semantic and phonological codes (Munding, Dubarry, & Alario, 2016; Strijkers & Costa 2016). Therefore, we tentatively argue for parallel mechanisms underpinning phonological processing regardless of the output modality. Moreover, the scalp distribution of the phonological effect observed in the present study is also consistent with previous research on spoken production in which a phonological effect emerged broadly distributed across posterior regions (e.g., Jescheniak, Schriefers, Garrett, & Friederici, 2002; Qu Damian, & Kazanina, 2012). Furthermore, the interval of 300-400 ms that hosts orthographic effects in the present study is consistent with findings from recent EEG studies on the investigation of word-form encoding underlying written production (e.g., Perret & Laganaro, 2012; Perret, Bonin, & Laganaro, 2014). It is worthy to note that the phonological effect is longer in duration (spanning roughly 300 ms) compared to the orthographic condition (roughly 100 ms). We speculate that longer duration of phonological effect may be associated with the complex relationship of orthographic and phonological systems in Chinese. Chinese is highly homophonic, with a spoken syllable corresponding on average to 11 Chinese characters. This one-to-many phonology-orthography mapping may cause the longer duration of phonological effects.

What could be the reason for the absence of a phonological effect in the behavioral data, in contrast to robust phonological effects found in previous studies with various tasks (e.g., Qu, Damian, Zhang, & Zhu, 2011; Qu, Damian, & Li, 2015)? The lack of visual feedback in our study could be a potential reason: In order to reduce artifacts of EEG data, participants were asked to keep looking at the screen while writing down their response. Without visual feedback, participants could not monitor their written execution and the position of holding the pen, and measurement of response latencies might not have been as reliable. However, Perret and Laganaro (2013) compared written production with and without visual feedback and found that both conditions involve the same central cognitive processes. Hence, the failure to obtain a phonological effect in our experiment is unlikely due to the lack of visual feedback. A more likely explanation centers on the “grain size” of phonological units to construct Chinese phonology. In the present study, phonological overlap was manipulated at the rhyme level. In Mandarin Chinese, syllables constitute the primary unit of phonological encoding while finer-grained units such as phonemes or rhymes play a comparatively weak during phonological encoding, and hence typically do not produce measurable effect in behavioral studies (e.g., Chen, Chen, & Dell, 2002; O’Seaghdha, Chen & Chen, 2010; Qu, Damian, & Kazanina, 2012). By contrast, the orthographic manipulation in the present study involved radicals, and numerous behavioral studies have documented radical-sized priming effects with Chinese individuals (e.g., Ding, Peng, & Taft, 2004). The radical-sized priming effect observed in the present study confirms the importance of this representational orthographic unit.

ERP waveforms showed that modulation of phonological overlap elicited a positive-going component with smaller ERPs elicited by the phonologically related than –unrelated condition. Modulation of orthographic relatedness elicited a negative-going component with the orthographically related condition generating smaller amplitude than the unrelated condition. In both cases, the word-form relatedness elicited smaller amplitudes, reflecting less demand. The polarity of orthographic and phonological ERP effects in the current study is compatible with those observed in studies of orthographic and phonological effects in the word recognition literature. Several studies consistently documented that P200 is sensitive to phonological processing. For example, Carreiras, Vergara and Barber (2005) found that a syllable manipulation produced an early ERP effect in P200 time window. Moreover, previous studies have suggested that orthographic manipulations could modulate a negative-going waveform characterized as N250 (Carreiras, Gillon-Dowens, Vergara, & Perea, 2009; Carreiras, Perea, Vergara, & Pollatsek, 2009; Dunabeitia, Molinaro, Laka, Estevez, Carreiras, 2009; Grainger & Holcomb, 2009; Grainger, Kiyonaga, & Holcomb, 2006). For instance, Grainger, Kiyonaga, and Holcomb (2006) observed that transposed-letter priming which reflects orthographic processing modulated the N250 component, with the orthographic manipulation eliciting smaller amplitude than the unrelated condition. In conjunction with these studies above, we speculate that in the current study, the phonological manipulation elicited P200, and the orthographic manipulated elicited N250.

Yet, more important is the information provided by EEG with regard to the relative time course of orthographic and phonological encoding. The main finding that activation of phonological codes evidently precedes access to orthographic codes extends on initial evidence from behavioral tasks (see introduction; Damian & Qu, 2013; Qu, Damian, Zhang, & Zhu, 2011; Zhang & Damian, 2010). This pattern has important consequences for current thinking about how orthographic output is generated. As outlined in the Introduction, at present the most plausible model of orthographic encoding postulates a direct link from meaning to orthography, as well as an indirect route from meaning to phonology to orthography (e.g., Bonin, Peereman, & Fayol, 2001). The current results support the notion that access to sound from meaning is efficient and speedy and has “priority”. Indeed, some independent evidence suggests that individuals, when presented with objects, involuntarily access their spoken names (Bles & Jansma, 2008). By contrast, access to orthographic information appears to be slower, probably due to the fact that this link is less relevant in everyday live, and has undergone less practice across the lifetime than retrieval of corresponding spoken codes. Hence both routes contribute to orthographic encoding (as evidenced by the independent behavioral priming effects obtained for orthographic and phonological overlap), but with different time courses. Our results highlight the potential of obtaining information about the time course of cognitive processes via EEG, something which is difficult or impossible to achieve via behavioral measures alone.

Our finding of earlier access to phonology than to orthography in written production is in agreement with our earlier results from purely behavioral results (e.g., Qu, Damian, Zhang, & Zhu, 2011) summarized in the Introduction. However, this pattern contrasts with both behavioural (Zhang & Wang, 2015) and EEG results (Zhang & Wang, 2016) in which relatively late phonological effects were found. Here we consider possible reasons for this discrepancy. In Zhang and Wang’s studies, all object names were monosyllabic, whereas in our experiments (including the present one), stimuli were disyllabic. A salient characteristic of Chinese phonology is its pervasive homophony: each spoken syllable maps onto 11 characters on average, and therefore a spoken word/syllable corresponding to a single character is typically ambiguous with regard to its meaning. By contrast, disyllabic homophones are relatively rare in Mandarin Chinese. We speculate that access to sound from meaning might be less efficient and speedy for monosyllabic compared to disyllabic words, due to the extensive activation of homophones in the preparation of the former but not the latter targets. We argue that disyllabic words as stimuli should be preferred due to their uncommon homophony and thus less ambiguity. Moreover, disyllabic words are the most prevalent type (63.87%) whereas monosyllabic words account for only 2.78% (Chinese Linguistic Data Consortium, 2003), hence disyllabic words constitute the more representative materials of Chinese. This account is admittedly speculative but generates clear predictions; future studies should vary the number of characters of response words to examine this possibility.

Another potentially relevant aspect arises from the fact that all previous studies which had explored contributions of phonology to handwriting (see Introduction) elicited responses consisting of single words. By contrast, in our study participants wrote short utterances consisting of two words. Processing in single word and phrase production is likely to differ in important aspects (see Bürki & Laganaro, 2014, for a recent exploration of this contrast in spoken production). A further possibility worth considering is that when the task requires the coordination of multiple constituents (as in our experiment), the role of phonology in writing might be particularly pronounced, due to the necessary buffering of lexical constituents. The specific need for such buffering presumably arises from the slow execution speed of handwritten output, compared to speaking. As rehearsal in short term memory is generally assumed to involve phonological codes, effects of phonology in writing might emerge more prominently (and perhaps earlier) in multi- than in single-word response tasks (see Tainturier & Rapp, 2001, pp. 267-268). However, we argue that on balance, a task such as ours which requires the output of multiple words should be preferred because it resembles writing outside the laboratory more than tasks do which require the written production of single, isolated words. Moreover, it is worth considering the scope of advance planning of word-form properties when participants write down multiple constituents (as in our experiment). The effects of word-form manipulations between prenominal color adjectives and object nouns reported in the present study suggest that writers planned the entire phrase at the word-form level before initiating a response.

It is important to note that information regarding the time course of two cognitive processes is instructive only to the extent that the degree of overlap is comparable across the two kinds of processes. In our study, we manipulated orthographic relatedness via a radical shared between color and object name, and phonological overlap via a shared rhyme. From a linguistic point of view, there is good reason to believe that the two types of overlap are comparable. Chinese characters, as basic orthographic units, map onto spoken syllables. Radicals are the largest sub-lexical orthographic units, while rhymes are the largest phonological units. Hence, experimental manipulations comparing radical and rhyme manipulations in Chinese involve a similar representational “grain size”, which confirms that it is legitimate to compare the two experimental conditions concerning their respective time courses in EEG.

Finally, it is worth highlighting that the position of overlap within the response words differed for orthographic and phonological overlap: the shared radicals occupied word-initial position, whereas the overlapping rhymes occupied non-initial positions. This was the case due to constraints on stimulus selection; however, could it be that the difference in relative time course for the two types of form overlap arose due to the overlapping portions occupying different positions within the response words? We think this is unlikely to be relevant: given the sequential nature of production, if position of overlap was relevant, (word-initial) orthographic effects in our study should have appeared *earlier* than (non-initial) phonological effects; however, EEG results showed the opposite pattern. Nonetheless, future studies should attempt to exclude this potential confound.

In sum, our study shows that phonological codes constrain written word preparation at a relatively early point in time, starting approximately 206 ms post picture onset. Access to orthographic codes is slower, starting at approximately 298 ms post picture onset. These findings will provide important constrains on psycholinguistic models of handwritten production.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (NSFC), No. 31771212, Youth Innovation Promotion Association CAS, and the German Research Foundation (DFG) and the NSFC in project Crossmodal Learning, DFG TRR-169/NSFC No. 61621136008 to the first author.

References

Afonso, O., & Álvarez, C. J. (2011). Phonological effects in handwriting production: Evidence from the implicit priming paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 1474-1483.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390-412.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68,* 255–278.

Bates, D., & Maechler, M. (2009). lme4: Linear mixed-effects models using S4 classes (R package Version 0.999375-32) [Computer software]. Retrieved from http://CRAN.R-project.org/package?lme4

Baus, C., Strijkers, K., & Costa, A. (2013). When does word frequency influence written production? *Frontiers in Psychology, 4*(December), 963. doi:10.3389/fpsyg.2013.00963

Bles, M., & Jansma, B. M. (2008). Phonological processing of ignored distractor pictures, an fMRI investigation. *BMC Neuroscience, 9,* 20-34.

Bonin, P., Fayol, M., & Peereman, R. (1998). Masked form priming in writing words from pictures: Evidence for direct retrieval of orthographic codes. *Acta Psychologica*, *99,* 311-328.

Bonin, P., Peereman, R., & Fayol, M. (2001). Do phonological codes constrain the selection of orthographic codes in written picture naming? *Journal of Memory and Language, 45,* 688-720.

Bub, D., & Kertesz, A. (1982). Evidence for lexicographic processing in a patient with preserved written over oral single word naming. *Brain, 105,* 697-717.

Bürki, A., & Laganaro, M. (2014). Tracking the time course of multi-word noun phrase production with ERPs or on when (and why) cat is faster than the big cat. *Frontiers in Psychology, 5*, 586.

Carreiras, M., Gillon-Dowens, M.,Vergara, M.,& Perea, M. (2009). Are vowels and consonants processed differently? ERP evidence with a delayed letter paradigm. *Journal of Cognitive Neuroscience, 21,* 275–288.

Carreiras, M., Perea, M., Vergara, M., & Pollatsek, A. (2009). The time course of orthography and phonology: ERP correlates of masked priming effects in Spanish. *Psychophysiology*, *46*(5), 1113–22.

Carreiras, M., Vergara, M., & Barber, H. (2005). Early event-related potential effects of syllabic processing during visual word recognition. *Journal of Cognitive Neuroscience, 17*(11), 1803–1817.

Chen, J.-Y., Chen, T.-M., & Dell, G. S. (2002). Word-form encoding in Mandarin Chinese as assessed by the implicit priming task. *Journal of Memory and Language, 46,* 751–781.

Chinese Linguistic Data Consortium. (2003). 现代汉语通用词表[Chinese lexicon] (CLDC-LAC-2003-001). Beijing, China: Tsinghua University, State Key Laboratory of Intelligent Technology and Systems, and Chinese Academy of Sciences, Institute of Automation.

Costa, A., Strijkers, K., Martin, C., & Thierry, G. (2009). The time course of word retrieval revealed by event related brain potentials during overt speech. *Proceedings of the National Academy of Sciences of the United States of America, 106,* 21442-21446.

Damian, M. F., & Dumay, N. (2007). Time pressure and phonological advance planning in spoken production. *Journal of Memory and Language, 57,* 195-209.

Damian, M. F., & Dumay, N. (2009). Exploring phonological encoding through repeated segments. *Language and Cognitive Processes, 24*, 685-712.

Damian, M. F., & Qu, Q. (2013). Is handwriting constrained by phonology? Evidence from Stroop tasks with written responses and Chinese characters. *Frontiers in Psychology,* 4(October), 765. doi:10.3389/fpsyg.2013.00765

Damian, M. F., & Stadthagen-Gonzalez, H. (2009). Advance planning of form properties in the written production of single and multiple words. *Language and Cognitive Processes, 24,* 555-579.

Damian, M. F., Dorjee, D., & Stadthagen-Gonzalez, H. (2011). Long-term repetition priming in spoken and written word production: Evidence for a contribution of phonology to handwriting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 813-826.

Ding, G., Peng, D., & Taft, M. (2004). The nature of the mental representation of radicals in Chinese: a priming study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(2), 530–539.

Dunabeitia, J.A., Molinaro, N., Laka, I., Estevez, A., & Carreiras, M. (2009). N250 effects for letter transpositions depend on lexicality: 'casual' or 'causal'? *Cognitive Neurosciences and Neuropsychology, 20,* 381-387.

Ganushchak, L. Y., Christoffels, I. K., & Schiller, N. O. (2011). The use of electroencephalography in language production research: a review. *Frontiers in Psychology,* 2(September), 208. doi:10.3389/fpsyg.2011.00208

Geschwind, N. (1969). Problems in the anatomical understanding of the aphasias. In A. L. Benton (Ed.), *Contributions to clinical neuropsychology* (pp. 107–128). Chicago, IL: Aldine.

Grainger, J., & Holcomb, P.J. (2009). An ERP investigation of orthographic priming with relative-position and absolute-position primes. *Brain Research, 1270,* 45-53.

Grainger, J., Kiyonaga, K., & Holcomb, P. J. (2006). The time-course of orthographic and phonological code activation. *Psychological Science, 17,* 1021–1026.

Guthrie, D., & Buchwald, J. S. (1991). Significance testing of difference potentials. *Psychophysiology, 28,* 240–244.

Indefrey, P. (2011). The Spatial and Temporal Signatures of Word Production Components: A Critical Update. *Frontiers in Psychology,* 2(October), 1–16. doi:10.3389/fpsyg.2011.00255

Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition, 92*,101-144.

Jescheniak, J. D., Schriefers, H., Garrett, M. F., & Friederici, A. D. (2002). Exploring the activation of semantic and phonological codes during speech planning with event related potentials. *Journal of Cognitive Neuroscience, 14*, 951-964.

Laganaro, M., & Perret, C. (2011). Comparing electrophysiological correlates of word production in immediate and delayed naming through the analysis of word age of acquisition effects. *Brain Topography, 24,* 19-29.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 1–75.

Liu, Y., Hao, M., Li, P., & Shu, H. (2011). Timed Picture Naming Norms for Mandarin Chinese. *PLoS ONE, 6*(1): e16505. doi:10.1371/journal.pone.0016505

Lupker, S. J., & Katz, A. N. (1981). Input, decision, and response factors in picture–word interference. *Journal of Experimental Psychology: Human Learning and Memory, 7*, 269-282.

Luria, A. R. (1970). *Traumatic aphasia: Its syndromes, psychology, and treatment.* The Hague, The Netherlands: Mouton.

Miozzo, M., Pulvermüller, F., & Hauk, O. (2014). Early parallel activation of semantics and phonology in picture naming: evidence from a multiple linear regression MEG study. *Cerebral Cortex*. 253343–3345, http://dxdoi.org/10.1093/cercor/bhu137.

Munding, D., Dubarry, A. S., & Alario, F. X. (2016). On the cortical dynamics of word production: A review of the MEG evidence. *Language, Cognition and Neuroscience, 31*, 441-462.

O'Seaghdha, P. G., Chen, J. Y., & Chen, T. M. (2010). Proximate units in word production: Phonological encoding begins with syllables in Mandarin Chinese but with segments in English. *Cognition, 115,* 282-302.

Perret, C., & Laganaro, M. (2012). Comparison of Electrophysiological Correlates of Writing and Speaking: A Topographic ERP Analysis. *Brain Topography, 25,* 64-72.

Perret, C., & Laganaro, M. (2013). Why are written picture naming latencies (not) longer than spoken naming? *Reading and Writing, 26*, 225-239.

Perret, C., Bonin, P., & Laganaro, M. (2014). Exploring the multiple-level hypothesis of AoA effects in spoken and written object naming using a topographic ERP analysis. *Brain and Language, 135,* 20–31.

Qu, Q. Q., Damian, M. F., & Kazanina, N. (2012). Sound-sizes segments are significant for Mandarin speakers. *Proceedings of the National Academy of Sciences of the United States of America, 109,* 14265-14270.

Qu, Q. Q., Damian, M. F., & Li, X. (2015). Phonology contributes to writing: Evidence from a masked priming task. *Language, Cognition & Neuroscience.* DOI: 10.1080/23273798.2015.1091086

Qu, Q. Q., Damian, M. F., Zhang, Q. F., & Zhu, X. B. (2011). Phonology contributes to writing: Evidence from written word production in a nonalphabetic script. *Psychological Science, 22*, 1107-1112.

Qu, Q. Q., Zhang, Q. F., & Damian, M. F. (2016). Tracking the Time Course of Lexical Access in Orthographic Production: An Event-Related Potential Study of Word Frequency Effects in Written Picture Naming. *Brain and Language*, *159*, 118-126.

Rapp, B., Benzing, L., & Caramazza, A. (1997). The autonomy of lexical orthography. *Cognitive Neuropsychology, 14,* 71-104.

Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture–word interference studies. *Journal of Memory and Language, 29,* 86–102.

Scinto, L. F. (1986). *Written language and psychological development*. New York: Academic Press.

Shen, X. R., Damian, M. F., & Stadthagen-Gonzalez, H. (2013). Abstract graphemic representations support preparation of handwritten responses. *Journal of Memory and Language, 68,* 69-84.

Snodgrass, J. C., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for names agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 174-215.

Strijkers, K., & Costa, A. (2011). Riding the lexical speedway: a critical review on the time course of lexical selection in speech production. *Frontiers in Psychology,* 2(December), 356. doi:10.3389/fpsyg.2011.00356

Strijkers, K., Costa, A., & Pulvermüller, F. (2017). The cortical dynamics of speaking : Lexical and phonological knowledge simultaneously recruit the frontal and temporal cortex within 200 ms. *NeuroImage*, *163*, 206–219.

Strijkers, K., Costa, A., & Thierry, G. (2010). Tracking lexical access in speech production: electrophysiological correlates of word frequency and cognate effects. *Cerebral Cortex, 20*, 913-928.

Strijkers, K., Holcomb, P. J., & Costa, A. (2011). Conscious intention to speak proactively facilitates lexical access during overt object naming. *Journal of Memory and Language, 65*(4), 345–362.

Tainturier, M.-J., & Rapp, B. (2001). The spelling process. In: B. Rapp (Ed.), *The Handbook of Cognitive Neuropsychology.* Taylor & Francis.

Thierry, G., Cardebat, D., & Demonet, J. F. (2003). Electrophysiological comparison of grammatical processing and semantic processing of single spoken nouns. *Brain Research, 17*, 535-547.

Wang, C., & Zhang, Q. (2015). Phonological codes constrain output of orthographic codes via sublexical and lexical routes in Chinese written production. *PloS One*, 10(4). http://dx.doi.org/10.1371/journal.pone.0124470

Zhang, Q., & Damian, M. (2010). Impact of phonology on the generation of handwritten responses: Evidence from picture-word interference tasks. *Memory & Cognition, 38,* 519-528.

Zhang, Q., & Wang, C. (2015). Phonology is not accessed earlier than orthography in Chinese written production: evidence for the orthography autonomy hypothesis. *Frontiers in Psychology,* 6(April), 448. doi:10.3389/fpsyg.2015.00448

Zhang, Q., & Wang, C. (2016). The Temporal Courses of Phonological and Orthographic Encoding in Handwritten Production in Chinese: An ERP Study. *Frontiers in Human Neuroscience,* 10:417. doi: 10.3389/fnhum.2016.00417

Zhang, Q., & Weekes, B. S. (2009). Orthographic facilitation effects on spoken word production: Evidence from Chinese. *Language and Cognitive Processes, 24*, 1082–1096.

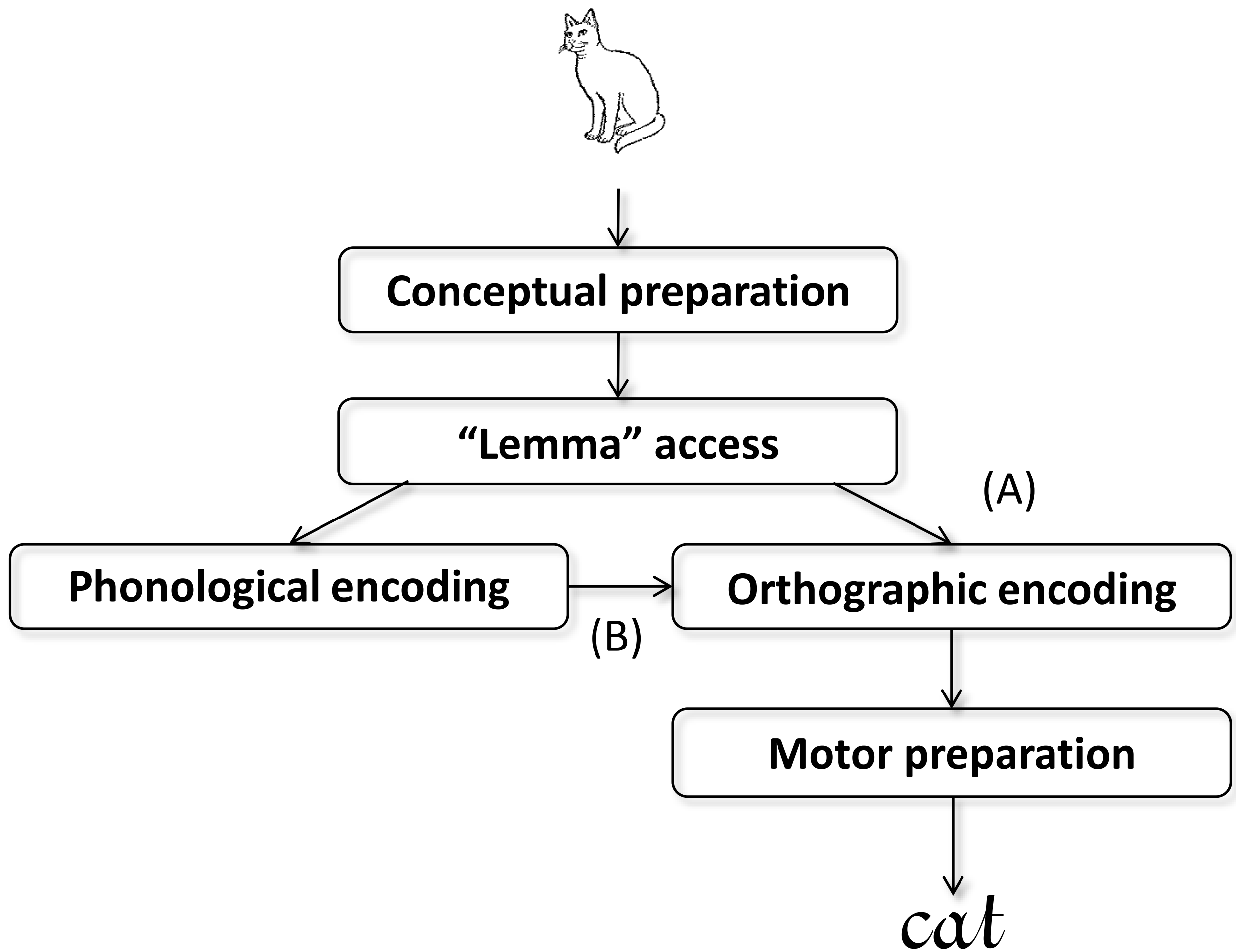
Zhao, H., La Heij, W., & Schiller, N. O. (2012). Orthographic and phonological facilitation in speech production: New evidence from picture naming in Chinese. *Acta Psychologica, 139*(2), 272–280.

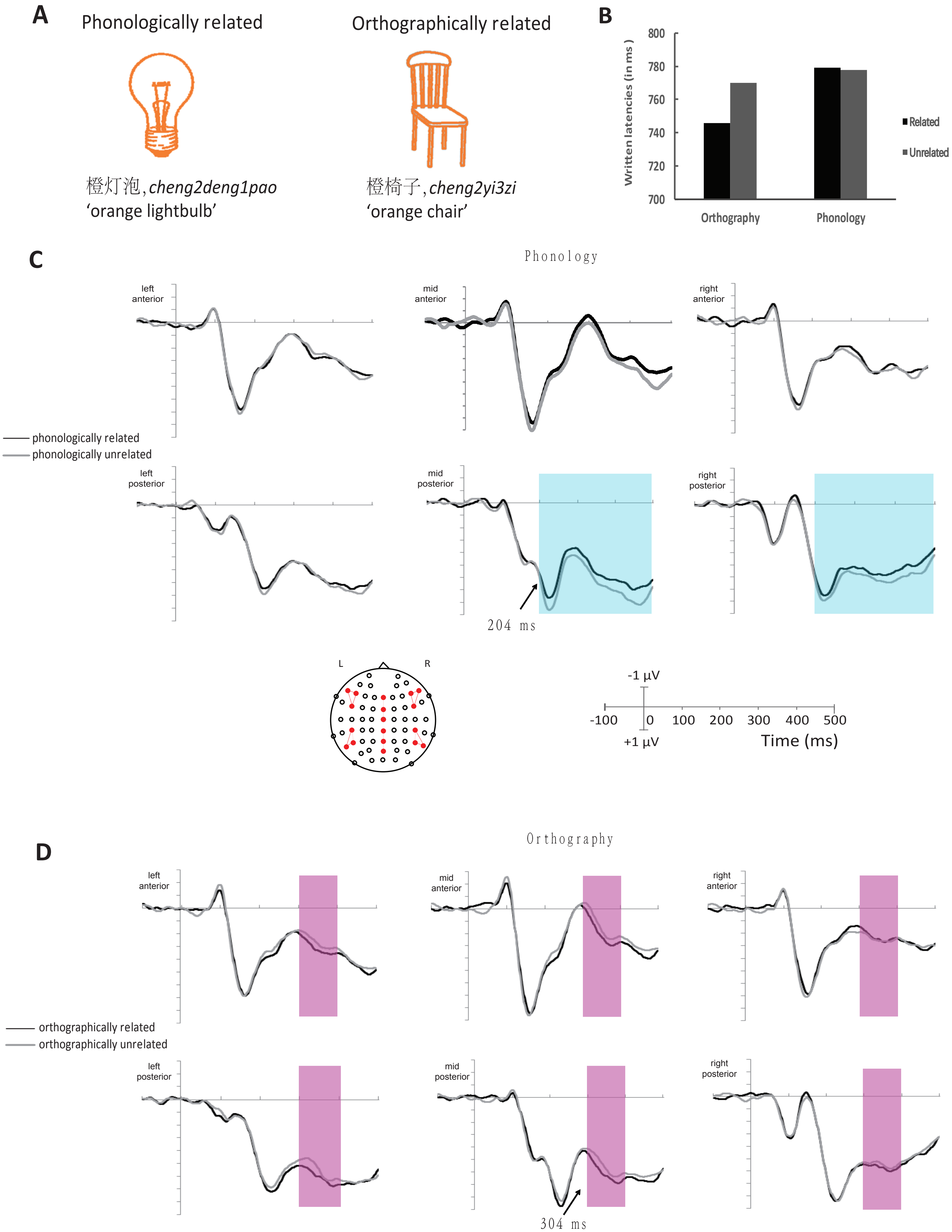
Appendix. Stimuli Used in the Experiment

Color name	Condition			
	Phonologically-related	**Phonologically-unrelated**	**Orthographically-related**	**Orthographically -unrelated**
红(red, hong2)	松鼠(squirrel, song1shu3)	蛋糕(cake, dan4gao)	线轴 (cotton reel, xian4zhou2)	杯子 (cup, bei1zi)
红(red, hong2)	公路(road, gong1lu4)	风筝(kite, feng1zheng)	绵羊 (sheep, mian2yang2)	梳子 (comb, shu1zi)
红(red, hong2)	钟(clock, zhong1)	灯泡(lightbulb, deng1pao4)	纺车 (spinning wheel,	枕头 (pillow, zhen3tou2)
蓝 (blue, lan2)	帆船(sailboat, fan1chuan2)	松鼠(squirrel, song1shu3)	花生 (peanut, hua1sheng1)	橡皮 (rubber, xiang4pi2)
蓝 (blue, lan2)	盘子(plate, pan2zi)	公路(road, gong1lu4)	花瓶 (vase, hua1ping2)	椅子 (chair, yi3zi)
蓝 (blue, lan2)	蛋糕(cake, dan4gao)	钟(clock, zhong1)	苍蝇 (fly, cang1ying)	柜子 (cupboard, gui4zi)
橙 (orange, cheng2)	绳子(rope, sheng2zi)	孔雀(peacock, kong3que4)	杯子 (cup, bei1zi)	花生 (peanut, hua1sheng1)
橙 (orange, cheng2)	风筝(kite, feng1zheng)	恐龙(dinosaur, kong3long2)	梳子 (comb, shu1zi)	花瓶 (vase, hua1ping2)
橙 (orange, cheng2)	灯泡(lightbulb,	公鸡(rooster, gong1ji1)	椅子 (chair, yi3zi)	苍蝇 (fly, cang1ying)
棕 (brown, zong1)	孔雀(peacock, kong3que4)	帆船(sailboat, fan1chuan2)	橡皮 (rubber, xiang4pi2)	线轴 (cotton reel, xian4zhou2)
棕 (brown, zong1)	恐龙(dinosaur,	盘子(plate, pan2zi)	枕头 (pillow, zhen3tou2)	绵羊 (sheep, mian2yang2)
棕 (brown, zong1)	公鸡(rooster, gong1ji1)	绳子(rope, cheng2zi)	柜子 (cupboard, gui4zi)	纺车 (spinning wheel, fang3che1)

Figure Caption

Figure 2. (A) Colored objects were presented to participants who wrote down color and picture name as an adjective-noun phrase. In the phonologically related condition, color and object names shared a rhyme; in the orthographically related condition, both shared a radical. Numbers indicate tone for each character; a neutral tone is not indicated. (B) Behavioral data show a non-significant phonological facilitation effect (-1 ms) and a significant orthographic facilitation effect (17 ms). (C) Grand average ERPs from 26 Chinese participants for phonological related (black line) and unrelated (grey line) conditions at six regions of interest (ROIs, shown as red filled circles on the electrode layout on the middle): left-anterior (electrodes: F7, F5, FC5), mid-anterior (Fz, FCz, Cz), right-anterior (F6, F8, FC6), left-posterior (P7, P5, CP5), mid-posterior (CPz, Pz, POz) and right-posterior (CP6, P8, P6). 0 ms represents the onset of a picture. In the posterior regions, the phonological related condition was significantly less positive in the 200- to 500-ms time interval, starting at 206 ms after picture onset (blue shading). (D) Grand average ERPs from 26 Chinese participants for orthographically related (black line) and unrelated (grey line) conditions at six regions of interest. In the left-middle regions, the orthographically related condition was significantly less negative in the 300- to 400-ms time interval, starting at 298 ms after picture onset (purple shading).


Conceptual preparation
"Lemma" access
(A)
Phonological encoding
Orthographic encoding
(B)
Motor preparation
cat



A
Phonologically related
Orthographically related
橙灯泡, cheng2deng1pao
'orange lightbulb'
橙椅子, cheng2yi3zi
'orange chair'
B
Written latencies (in ms)
800
780
760
740
720
700
Orthography
Phonology
Related
Unrelated
C
Phonology
left anterior
mid anterior
right anterior
phonologically related
phonologically unrelated
left posterior
mid posterior
right posterior
204 ms
L
R
-1 µV
-100
0
100
200
300
400
500
+1 µV
Time (ms)
D
Orthography
left anterior
mid anterior
right anterior
orthographically related
orthographically unrelated
left posterior
mid posterior
right posterior
304 ms